[illegible]

[illegible]

[illegible]

(8276323)

[redacted]

[redacted]
Department of [redacted]

University of Tulsa

[redacted]

[redacted]

[redacted]

[illegible]

Department of English
University of Turku

# EXPECTING THE DISCOURSE TOPIC IN NATIVE AND FOREIGN LANGUAGE TEXT COMPREHENSION

## 1. Introduction

### *1.1. The discourse topic*

Van Dijk (1977, 1980) defines the discourse topic as a macroproposition. This proposition is arrived at linguistically by applying a set of recursive macrorules to the surface representation of the text. The rules delete, reorganise and reformulate information in such a way that the resulting statement expresses, on a more global and abstract level, the essential information from the original discourse. For example, consider the short passage in (1).

(1) (a) As is now commonly known, the world's reserves of oil are dwindling fast, and some people predict serious energy crises in the future. (b) In consequence, alternative sources of energy are being investigated everywhere. (c) Wood and coal are now used increasingly in both commercial and home heating. (d) The use of wind and solar energy in producing electricity is no longer a curiosity, and cars using alcohol instead of gasoline are already being developed.

Macropropositional reduction is achieved by such rules as WEAK DELETION, which omits the information contained in the first clause as irrelevant; STRONG DELETION, which omits the idea of predicted energy crises as information that, although locally relevant, is by automatic inference relatable to the contents of the second clause of sentence (a); ZERO, which preserves the content of the second clause of (a) intact; and GENERALISATION, which creates a general statement about increasing interest in other sources of

energy from the topical sentence (b) and the following sentences which elaborate on it. The result of the application of these rules is two basic macropropositions, expressed informally as

(2) a. The world's reserves of oil are dwindling fast.
b. Alternative sources of energy are being taken up.

A CONSTRUCT rule may operate on the two macropropositions derived in the first cycle, resulting in a single, more general statement such as

(3) Alternative sources of energy are starting to replace oil.

Statement (3) is one possible way of expressing the global meaning of the discourse, and thus provides a usable summary. Its contents would also typically be represented in a title to the passage (e.g. 'Alternatives to oil'). It is such global propositions that represent discourse TOPICS in van Dijk's analysis. He thus equates the notion of discourse topic with a statement of gist.

In this paper a narrower definition of discourse topic will be used. The topical macroproposition in van Dijk's sense is "about" a notion, and gives a piece of new information about this notion. Thus it divides into a global theme and rheme. In the following, the discourse topic is understood as the notion to which the theme of the topical macroproposition refers (the referent of *alternative sources of energy* in the passage above). This definition is inadequate in practice to describe the topic in complex texts. For example, passage (1) above is not merely about alternative sources of energy, but instead about the fact that alternative sources of energy are replacing oil. However, the simpler definition will be adequate for some cases, e.g. descriptive texts where the properties of an object or person are enumerated. The materials of the present study are typically of the kind where the dominant text strategy is based on the iteration of a single theme (cf. Enkvist 1973, 1975) and the presentation of various bits of new information about the topical referent.

To summarise, we will say here that, for the materials to be used in the following, the macropropositional theme is taken as the discourse topic. It represents the point of departure for the discourse, and provides the answer

to the reader's question of immediate concern (Keenan and Schieffelin 1975): "What (notion) is the writer giving me information about?".

### *1.2. Signals for the discourse topic*

Various signalling devices are available to the writer for indicating how the information is "staged" (Grimes 1975, Clements 1979) in the discourse, i.e., which element of the content has been "thematised" (cf. Perfetti and Goldman 1974, 1975) as the referential core for the text. If the topical signals are properly formulated, and are interpreted correctly by the reader, the thematised notion becomes "foregrounded", i.e. placed in sharp focus in the mind of the reader (cf. Chafe 1972; Sanford & Garrod 1981).

The selection of a particular element of meaning as the central point of departure is necessary for the discourse to appear coherent. An extreme form of theme progression (Enkvist 1973) where the rheme of a sentence is turned into the theme of the following sentence, about which a piece of new information is again provided, yields text where linguistic cohesion may be faultless but where a coherent discourse topic never emerges. A more considerate text strategy is one which allows the selected topic to prevail for a certain stretch of text through theme iteration. One consequence of this is that expressions pointing to the discourse-topical referent will appear as themes in several of the sentences of the discourse, so that sentence themes function as cues to the topic of the discourse. The chief linguistic mechanism of maintaining topical coherence and signalling topical continuity is thus reference, e.g. explicit repetition of the topic or the use of other coreferential expressions. The cues may also be more indirect, as when reference relations are expressed through hyponyms and hyperonyms, or merely implied on the basis of shared experience. The perceived topic invokes a schema which is used by the reader to make the necessary inferences.

In many cases, a direct cue to the discourse topic is the topical sentence, the central sentence of a paragraph. Normally it appears paragraph-initially, sometimes after an introductory sentence (as in the case of (b) in passage (1) above), or paragraph-finally. Since it is the base to which further items of information in other sentences are tied by reference, it tends to be an autosemantic, often generic statement, and usually contains concepts that are

hyperonymic in relation to the key concepts in the sentences that are below it in the macrostructural hierarchy.

The statement containing the discourse topic may also be identified through rhetorical, illocutionary signals of sentence function. Sentence (b) in (1) above is an example of a topically central general statement that can be identified as such through the presence of clauses or sentences that specify its contents by examples or explanations. The sentence themes of the specifying statements naturally tend to be hyponymic in relation to the theme of the topical sentence.

The cues to discourse topics also include non-topical organising material (Lautamatti 1980) used to structure the text and to signal the writer's perspective. Perhaps the clearest example is straightforward metatext which points out the element of content selected as the topic ('I will next deal with the search for alternative sources of energy'). Related to such structuring statements are Meyr's (1975) 'explicit rhetorical predicates', i.e. expressions which indicate schematic rhetorical divisions of text into functional chunks such as setting, problem, solution, etc. ('A burning problem in today's world is finding alternatives to oil.'), which have obvious topical significance.

Discourse-topical information may also be signalled by the presence of text-internal interim summaries, which perform a similar macropropositional function as topical sentences, but with larger stretches of discourse. The title, already mentioned, and typographical conventions such as paragraph separation also serve as cues to the topic and the extent of its span.

### *1.3. Processing implications*

The understanding of continuous discourse involves an interaction between data driven and conceptually driven processing. The input is subjected to data-dependent recodings through which a meaning representation is generated from the bottom up. Simultaneously, the string of language, on which the analysis has just begun, makes contact with the knowledge structures in the reader's mind. This knowledge is utilised by processes which generate top-down expectations concerning microlevel (word, phrase, clause, sentence) structure as well as more global features (central content, schematic structure of the discourse, properties of objects and events in the real world, etc.).

The predictive hypotheses supplement the data-driven analysis, and involve a precoding of information, i.e., doing part of the processing in advance to facilitate the on-line handling of the input string. Data-driven processing is necessary for the generation and checking of the knowledge-based predictions: top-down processing must naturally be based on the data at hand. Conceptually driven processing is similarly likely to be present even in the earliest phases of comprehension after the purely perceptual operations. Its function is to direct the data-driven analysis while simultaneously being constrained by the bottom-up processes. Thus, what is sometimes seen as a strict dichotomy between the two modes is in reality a two-way interaction. The notion of interactive processing is inherent in comprehension models that rely on the notion of expectation but not necessarily in serial models.

The question of interest is what the interaction of these processing modes is like; for example, in what way the characteristics of the comprehender, input material, or situation determine the extent to which reading comprehension is facilitated by expectations concerning such global features as the discourse topic.

The processing implications of the discourse topic centre around the concept of the limited-capacity working memory. In well-formed text with sufficient topical continuity, the function of the discourse topic is to make efficient integration of new information possible by enabling the generation of coherence-based expectations. The topic guides and limits the memory search whereby an attempt is made to map the given information in the current sentence to antecedent information already stored (cf. Clark & Haviland 1977; Lesgold et al. 1979). If the theme of the current sentence corresponds to the foregrounded discourse-topical referent, the search is facilitated, and the new information in the sentence can be rapidly integrated. If the current sentence is about a notion no longer foregrounded, reinstatement of the backgrounded concept from long-term memory is necessary, which complicates processing, as does the situation where bridging inferences are required. If the memory search for antecedent information yields no results, the new information given in the sentence cannot be immediately integrated, in which case the theme of the current sentence is tentatively adopted as a new, additional discourse topic on a dynamically changing stack of topical hypotheses maintained by the comprehender.

In summary, the discourse topic is a macrostructural feature whose identification is likely to be a factor in the fluency of comprehension. Processing load may be reduced if the reader begins handling the current sentence in terms of the global topic. The way texts are structured on this variable, and the sensitivity of the reader to discourse-topical signals, should thus have clear processing consequences.

## 2. First and foreign language research on discourse topic processing

The recent interest in discourse phenomena in L1 comprehension is apparent in the large literature on the processing of the discourse topic and related textual constructs. In a sense, this literature follows the traditional line of psycholinguistic investigation: attempts are made to demonstrate the psychological reality of linguistic constructs by studying their processing consequences. As is well known, this line of investigation had its problems for syntax, and led to doubts about the value of linguistic categories and processes for the study of the processes of language use. On the other hand, it also led to scepticism about the value of processing evidence for argumentation in linguistics.

The search for the psycholinguistic reality of text structure, however, is well motivated. Texts are concrete units. The form they take is determined by performance factors such as ease of production and understanding, optimal structuring from the point of view of recall or desired effect, and so on. A text is a unit of language which takes a certain form because it is intended to lead to certain types of processing consequences. Thus, the connection of structure and process is more immediate for texts than it is for the system-sentences of language.

Typical parameters whose effects have been investigated include propositionally defined content structure, the organisation of texts into functional (rhetorical) units, and the effects of staging (e.g. Meyer 1975; Frederiksen 1977; Kintsch and van Dijk 1980; Thorndyke 1977). Some studies related to the staging variable and topic foregrounding are briefly mentioned in the following.

Anaphora studies such as Lesgold et al. (1979) and Purkiss (1980, reported in Sanford & Garrod 1981) demonstrate that the ease of referential mapping

of the given information in the current sentence to antecedent information in the discourse depends crucially on what is currently foregrounded as topical information. The limited-capacity working memory cannot hold available as possible antecedents everything that was mentioned at some time in the discourse. Things to be held available are selected, and one basis for selection is the reader's hypothesis about topicality. If the discourse-topical status of a referent becomes uncertain (as a result of intervening topic-changing sentences), it becomes difficult to locate that referent as the antecedent for expressions in the current sentence, even though the referent occupies a thematically prominent position as the subject of the antecedent sentence. The result is increased reading time for the current sentence. If the antecedent is an object and rheme in the antecedent sentence, the intervening sentences tend to have an even larger effect on the reading times of the target, suggesting that the rheme of the antecedent sentence is less resistant than its theme to the availability-diminishing effect of the intervening sentences. The reason is presumably that the theme assumes discourse-topical status during the processing of the antecedent sentence, and remains, to a certain extent, on stage, in spite of the intervening material. The probable explanation for the increased processing load in the case of a long referential span is reinstatement matching (cf. Lesgold et al. 1979), which is obviously more costly than immediate matching of the given information with antecedent information.

Kieras (1978) found that topical coherence and the use of a top-down convention of placing high-level information first in a paragraph effectively directs subjects' choice of what is understood as the discourse topic (as measured by a title formulation task). Recall scores were higher for top-down (deductive) paragraphs than for bottom-up (inductive) paragraphs, but sentence reading times did not differ in the two conditions if subjects read the passage for a recall test. Good paragraph structure produced a facilitatory effect in reading times when a relatively artificial method of increasing cognitive load was used (asking subjects to read material where sentences from several passages were interleaved). Kieras (1981a) suggests that the number of items on the topic stack is a measure of passage coherence, and a factor in reading time. Adding a new candidate item to the stack increases processing load, since there will be more items competing for the limited capacity of the system. Another observation by the same author (1981b)

is that title formulation is facilitated and overall passage reading times are shortened if the passage is staged around a single referent functioning as a clearly identifiable discourse topic.

The interplay of the discourse topic with sentence syntax and thematics is further demonstrated by Perfetti and Goldman (1974; cf. also 1975), who investigated the efficiency of a thematised referent (discourse topic) as a memory probe. When the referent of the subject of a paragraph-final SVO sentence had been the topic, the subject functioned better as a probe than did the object word. When the object had been the topic, the subject and object were equal as probes. Thus, when the theme of the target sentence continued the discourse topic, the combined effect of discourse-topical and sentential foregrounding kept the subject word strongly salient. The equality of the subject and object words as probes in the condition where the target sentence object (rheme) had been foregrounded in the preceding context indicates that the object had now acquired a degree of salience by virtue of its discourse-topical status while sentence thematics maintained the salience of the subject's referent.

The importance of topic knowledge for interpreting and recalling vague, ambiguous or richly metaphorical text is clear from studies like Dooling & Lachman (1971), and the effects of topic knowledge are here related to the expectation-oriented notions of frame, schema, script and scenario (see e.g. Tannen 1979; Sanford & Garrod 1981). Once established, the topic helps to activate a stereotypical body of knowledge, which makes predictable aspects of content available even before the input signal has been fully analysed.

Less research has been carried out on the effects of topics and related discourse variables on foreign-language processing, which probably reflects a basically sentence-based approach to non-native performance as well as an orientation that Enkvist (1980) terms structural as opposed to processual. However, introspective self-monitoring data (Cohen & Hosenfeld 1981; Hosenfeld 1977) seem to indicate that successful FL performers are consciously aware of employing global reading strategies such as attempting to hold central passage content in mind while reading, whereas microlevel operations emerge as central in the intro(retro)spections of less successful language learners. Instruction may sway reading strategies toward the global type (cf. Hosenfeld 1977); pedagogical suggestions are given by Godfrey (1977)

and Clarke & Silberstein (1977). Chihara et al. (1977) is a demonstration that foreign language proficiency interacts with a coherence variable of a relatively crude type: it appears to make more difference to advanced learners than to less advanced learners whether the discourse is well-formed or consists of scrambled sentences. Using both oral reading miscue analysis and various silent reading tasks, Cziko (1977; 1980) has demonstrated that there are proficiency-dependent differences in the extent to which non-native readers employ discoursal cues, and according to Jacobs (1981), awareness of rhetorical signals of connection and illocutionary function is an important determinant of foreign-language proficiency.

## 3. The present study

The research described above indicates that the discourse topic is a significant factor in global comprehension. The topic also appears to guide the more local processes at sentence level and below. Methodologically, the bulk of L1 processing research on the topic and other discourse variables has utilised successive techniques in various recall paradigms, but simultaneous measures of on-line performance, especially total sentence reading times, have also been employed to some extent. Clements (1979), however, calls for more research concerning the on-line effects of staging variables, and suggests the use of eye-movement methodology and probes to interrupt reading at selected points to study on-line performance. Britton et al. (e.g. 1978) employ a subsidiary task of monitoring for a nonverbal signal during the reading of long texts to estimate processing load (cf. also Britton et al. 1979, who failed to demonstrate the on-line effects of Meyer's paragraph height with this method).

While there is some interest in discourse factors among foreign-language teachers and test constructors, few studies seem to have investigated their effects from a psycholinguistic angle by studying the on-line comprehension performance of foreign language users. The present study attempts to do that by examining the on-line processing consequences of discourse topic identification and foregrounding in native English readers and two groups of non-native (Finnish) users of English.

The general question is to what extent the reader develops expectations about the discourse topic, and whether these predictions guide and facilitate the processing of information contained in the individual sentences. The study also poses the question whether language proficiency has any effect on the identification and utilisation of discourse-topical information during reading. The focus of interest will be the final sentence in short paragraphs, and the compatibility of the theme of this sentence with the discourse topic foregrounded in the preceding context. In one condition, the theme of the paragraph-final sentence represents a continuation of the foregrounded discourse topic. In another condition it is a concept which was mentioned in the initial sentence of the text, but which has subsequently been backgrounded. An on-line measure of processing load is taken at a point immediately following the theme of the final sentence, in order to see whether the processing of the theme is affected by discourse-topical expectations that have arisen while the preceding context was being read. The measure used involves reaction times in a subsidiary task, visual word monitoring, which makes it possible to focus attention on the crucial thematic point as it comes up for processing.

What was said above about the linguistic signalling of discourse topics and their processing implications suggests that, for fluent L1 readers, the marking of a referent as topic, and the topical continuity of text, should facilitate not only recall but also on-line text processing. For these readers, comprehension probably begins with an attempt to construct an immediate macrostructural hypothesis about what the text is dealing with. The reader may then approach each new sentence with the hypothesised discourse topic in mind, trying to establish the relation of the current sentence to the topic (cf. Clark and Haviland 1977). Unless the end of the topical span is indicated by concluding signals, the starting point for the interpretation of a new sentence is the expectation that its theme will continue to be compatible with the established general topic. The strength of the signalling of this compatibility may of course vary. A particularly strong form of signalling is the repetition of the discourse topic in the themes of the individual sentences, and for this reason such iteration cannot continue long without the text becoming stylistically unacceptable.

The assumption made here for fluent L1 readers is that processing load is relatively low at a point where the theme of the final sentence is being

not been foregrounded in the immediately preceding context, though it has been mentioned in the initial sentence, a reinstatement of this referent from long term memory is necessary. At a point immediately following the sentence theme, an increase in processing load, due to this reinstatement, should therefore be observed for the fluent reader.

However, the micro- and macrolevel processes of text understanding compete for the limited capacity available at a given moment. In view of this fact one may expect that the proportional share of capacity demanded by these processes can vary, for example as a function of the type of incoming material or reader characteristics such as cognitive style, previous training, or skill expertise. Lesgold and Perfetti (1978) suggest along these lines that poor general comprehension in native English text processing can in part be explained by slow letter and word scanning skills, or inefficient phonological coding. The implication is that poor comprehenders allocate much of their processing capacity to these lower-level operations. They are unable to pass smoothly over the local features of the message and to direct their spare capacity to such processes as topic or gist identification and the integration of information into a global model of discourse content.

To extrapolate from this to the foreign language situation, one may hypothesise that proficiency variations have an effect on where processing capacity goes in non-native reading. More specifically, it may be assumed that capacity allocation to the various levels of the message depends heavily on the automation of the microlevel processes (e.g. lexical access, parsing). The reason is that a certain amount of microlevel analysis is logically necessary for the construction and checking of top-down predictions before these can in turn facilitate microlevel processing. If the foreign language reader's local processes such as sentence comprehension function with a minimum of conscious control, capacity can be directed to the utilisation of macrolevel cues. If attention has to be centred on local processes, less capacity will be available for discourse operations such as constructing a gist representation from the sentences of the text, and predicting the thematics of incoming sentences from a hypothesis concerning the discourse topic.

If advanced non-native readers follow a similar strategy of topic expectation as native speakers, a facilitating effect of previous context should be

incompatibility, an inhibiting effect should be observed in much the same way as with native speakers. The fact that processing takes place in a foreign language may, however, introduce an overall increase in processing load even in subjects who are highly fluent. As a result, advanced non-native subjects should be slower than native speakers on a reaction time measure in the subsidiary task.

If less advanced non-native readers allocate a considerable portion of processing capacity to microlevel features of the language (morphological processes, basic syntactic relations) rather than global features, this should be reflected as a relatively small difference in processing ease for the paragraph-final sentence in the two conditions of topical compatibility and incompatibility. The assumption is thus that less advanced non-native readers may not always reach the depth of analysis required by adequate comprehension of macrostructure. Global discourse-level features may pass unnoticed, and even if they are perceived at some level, less advanced non-native readers may not use them to direct lower-level processing as efficiently as advanced non-native readers do. Again, if the measure of processing load is sensitive to foreign-language proficiency variations, the overall response times of less advanced foreign-language subjects in the subsidiary task should be slower than those of advanced non-native readers.

One would thus expect a global topic variable to interact with proficiency level in foreign language text comprehension. Advanced non-native readers should be sensitive to staging signals, and they should be able to recognise and keep in mind the discourse topic better than less advanced non-native readers. They should be able to construct topical expectations fairly efficiently during reading, and so their on-line comprehension performance should be affected by manipulations of topic structure to a greater degree than the performance of less advanced non-natives.

### *3.2. Subjects*

A group of thirty 2nd and 3rd-year students of English from the Department of English, University of Turku acted as an "Advanced" (ADV) group.

ceiving intensive language skills training during their two first years at university. A group of 28 students from the Faculty of Social Sciences of the same University acted as an "Intermediate" (INT) group. The latter group were enrolled in a course for students who had not yet reached a passing mark in an ESP reading comprehension test required by their Faculty. In addition, a group of 9 native speakers of English participated as a control group (NS).

### *3.3. Materials*

Six pairs of experimental passages and eight filler passages were constructed. Each passage had the following structure:

| | | |
|---|---|---|
| TITLE | | |
| CONTEXT SENTENCE 1<br>CONTEXT SENTENCE 2<br>CONTEXT SENTENCE 3 | } | approximately 75 words |
| TARGET SENTENCE | | 17 words (ASVX) |

Each pair of test passages was produced from a single original by staging the title and context section around two different referents. The function of the title was to give a particularly strong clue about the discourse topic. After the title, the first context sentence introduced two referents. The one in the rheme corresponded to the referent mentioned by the title. This structure of the initial sentence was used in order to achieve the effect of a reasonably natural introductory sentence. After a theme progression (Enkvist 1973), this referent appeared as the theme of the second and third context sentences, and was thus consolidated as the discourse topic, and foregrounded. The resulting two passages (A and B) now "were about" two different things, both of which, however, had been mentioned in the initial sentence.

The Condition A and Condition B contexts were both followed by an identical closing sentence, the target sentence. It began with an adverbial, which was followed by a subject NP, a verb, and a final constituent (object or complement). The subject NP (and theme) of this sentence either represented the continuation of the passage topic (Condition A), or reintroduced, by exact repetition or the use of a coreferential NP, the non-topical referent from the first context sentence (condition B). Table 1 gives a sample of the materials.

### *3.4. Procedure*

On-line processing during the target sentence was studied using a visual word-monitoring task. The subject was seated in front of a display terminal connected to a DEC20 computer and a microprocessor-controlled reaction time analyser (RTA). The presentation of the items was controlled by two programs, one in the terminal and the other in the DEC20, operating in synchrony and capable of presenting the items to subjects in real time. Reaction times were measured to an accuracy of 1 msec by the measurement and storage program of the RTA. The instrumentation and programs are described in Tommola et al. (1982; forthcoming), and Salmela and Tommola (1982).

The session began with 10-15 practice trials identical in form to the test items proper, after which the test items were presented. In each item, the title and the three context sentences were displayed for a constant reading time sufficient for the intermediate subjects. Instructions emphasised absorbing the main points of information from the text. At the end of the reading time, the screen was cleared, and the target sentence was presented, preceded by a 1-second display of the target word. The target sentence was typed onto the screen by the program at the rate of 10 characters per second, starting from a fixation point indicated on the screen. The entire sentence remained visible for 1 sec after the final letter had appeared. The task of the subject was to depress a microswitch upon seeing the pre-determined word, while simultaneously reading the target sentence for meaning.

The comprehension of the target sentences was controlled by asking the subjects to produce orally either paraphrase or a free translation of the

content. In addition, one comprehension question was presented on the context sentences. The purpose was to force the subjects to process the context at a deep enough level for possible topic effects to emerge, and similarly to ensure that they would read the target sentence for meaning.

Two sets of experimental materials were constructed. In each, half of the passages of interest appeared in condition A form, half appeared in condition B form. Each subject thus received 3 Version A passages and 3 Version B passages, together with 8 filler passages. Presentation order was separately randomised for both sets of material. Subjects were randomly assigned to either test material set.

The minimum and maximum limits for word monitor latencies to be included in analyses were pre-set at 100 msec and 1500 msec. Reaction times exceeding 1.5 seconds are generally thought to represent other than immediate comprehension processes in monitoring experiments.

*Table 1.* Sample of experimental materials

| VERSION A | VERSION B |
|---|---|
| TITLE:<br>Captain Archie Fry | TITLE:<br>The U.S. Coast Guard |
| CONTEXT SENTENCES:<br>Back in the 1930's the U.S. Coast Guard spent a lot of time trying to catch Captain Archie Fry, the chief organiser of illegal alcohol trade on the East Coast. | CONTEXT SENTENCES:<br>Back in the 1930's Captain Archie Fry and other organisers of illegal alcohol trade gave a lot of work to the U.S. Coast Guard on the East Coast. |
| Captain Fry, who was thoroughly familiar with the coastal waters, usually anchored his ship in a safe place in the dark of the night, and was then met by buyers who sailed out in small boats to purchase liquor. | The Coast Guard men, who were thoroughly familiar with the coastal waters, patrolled the likeliest trade routes in the dark of the night, searching for buyers who sailed out in small boats to purchase liquor. |

| | |
|---|---|
| The captain sailed only about two trips per year, but every time he sould thousands of gallons of illegal whiskey to underworld customers. | The Guard boats were on duty every single day of the year, and often confiscated thousands of gallons of illegal whiskey meant for underworld customers. |
| TARGET SENTENCE: For a number of years Fry *managed to avoid the probing searchlights of government ships. | TARGET SENTENCE: For a number of years Fry *managed to avoid the probing searchlights of government ships. |

In the target sentences, though not in the filler sentences, the word to be monitored for was always the 7th word in the sentence. In the table, the target word is preceded by an asterisk. Syntactically, it was a past-tense verb appearing after the subject NP (and theme) of the target sentence. The target verb was a relatively high-frequency item familiar to the foreign-language subjects. Furthermore, the verbs were judged to be general enough in meaning not to give the subject much indication of what the actual content of the target sentence would be. The target word was capable of occurring as the verb for the thematised referents of both Version A and Version B with roughly equal probability. The beginning of each target sentence was neutral with respect to both Version A and Version B contexts up to the subject NP head noun.

### *3.5. Results*

The subjects' oral paraphrases were evaluated by the experimenter for correct mention of the actor, action and goal of the target sentence. Three per cent of both native speaker and advanced non-native paraphrases were incomplete. In the intermediate group the percentage was 17. When a subject produced an incomplete paraphrase, the experimenter asked firstly what the target word had been and then asked the subject to talk about the whole text. In this free recall situation, the majority of subjects who had failed to produce a complete paraphrase were able to retrieve the central semantic

roles of the target sentence. It was concluded from this that incomplete paraphrases tended to result not so much from comprehension problems as from encoding difficulties associated with keeping the entire target sentence in mind while attempting to formulate it in different words or in the native language. The monitor latencies for these cases were included in the data for all groups.

All of the reaction times of the native speaker subjects fell within the pre-set limits of 100-1500 msec. The percentage of RTs discarded as overlong was 2% in both the advanced and the intermediate non-native group.

Occasionally subjects failed altogether to react to the target word. This happened more often with native English readers than non-natives. The most common reason given by subjects was involvement in the meaning of the passage and the consequent failure to pay attention to the target word.

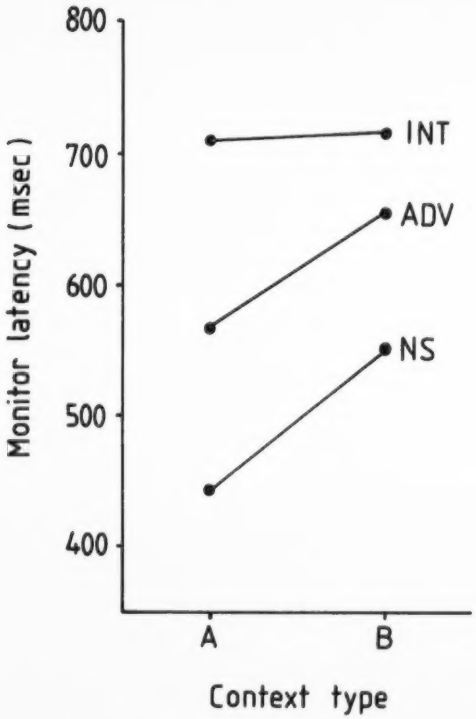
The total percentages of missing reaction times (overlong RTs and failures to react) in the three groups were 9% (intermediate), 5% (advanced), and 9% (native speakers).

*Monitor latencies.* The A and B monitor latencies were averaged for each subject, and an analysis of variance was performed on logarithmic transformations of the means, with context type (A, B) as a within-subjects factor, and proficiency (INT, ADV, NS) as a between-subjects factor. The pattern of the monitor latencies is given in Figure 1.

Figure 1 indicates, firstly, that there are distinct differences between the three proficiency-level groups in their overall word-monitor latencies. Advanced non-natives are, on the whole, approximately 100 msec faster than the intermediate group, and the native speaker group is approximately 130 msec faster than the advanced non-native group. The main effect of proficiency is significant, $F(2,67) = 21.26$, $p < 0.001$.

A significant main effect also emerged for the topic factor, $F(1,67) = 18.91$, $p < 0.001$. The results support the hypothesis that, on the whole, continued foregrounding of the discourse-topical referent facilitates ongoing processing at the point selected for observation here. When the non-foregrounded referent appears as the target sentence theme (condition B), word monitor latencies to an item immediately following the subject (theme) NP tend to be higher.

*Fig. 1.* Mean word monitor latencies for target sentence verbs in two context types and three groups of subjects. A = discourse topic as target sentence theme. B = a non-foregrounded referent as target sentence theme. INT = intermediate non-native, ADV = advanced non-native, NS = native speakers.)




The average word-monitor reaction time for native speakers in condition A was 115 msec shorter than the latency in condition B. The A-B difference in the advanced non-native group was 80 msec in the same direction. In the intermediate non-native group, the mean for condition A was 10 msec shorter than the condition B mean. The interaction between proficiency level and topic is significant, $F(2,67) = 3.86$, $p = 0.03$. Thus, while the native

speaker and advanced non-native groups are responding almost identically to the manipulation of the topic variable, the intermediate group displays a different pattern by not being similarly affected.

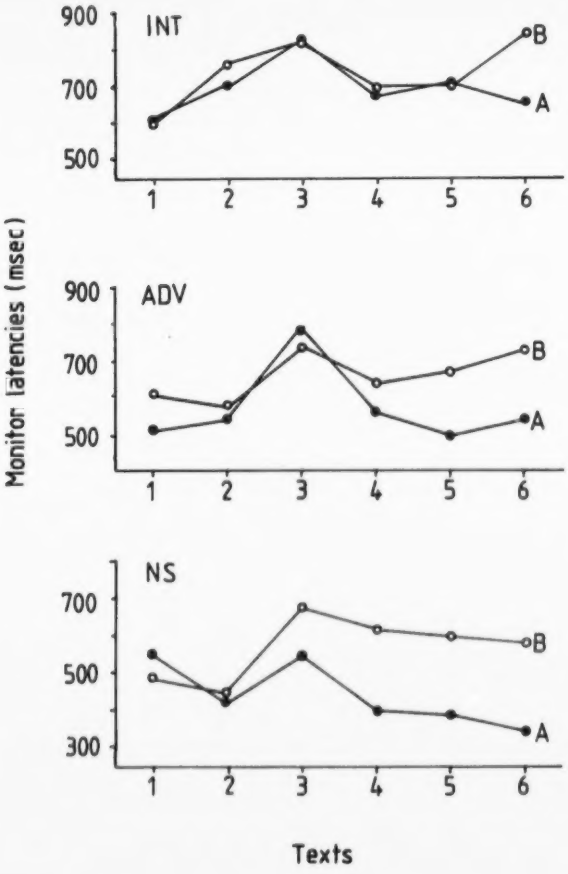
The mean A latency was shorter than the mean B latency for all native speaker subjects. In the advanced non-native group, 3 out of the 30 subjects had a mean A latency that was over 10 msec longer than the mean B reaction time. Fifteen out of the 28 subjects in the intermediate non-native group displayed the latter pattern. Thus, the native speaker and advanced groups appear to be homogeneous with respect to the effects of the topic variable, whereas the performance of the intermediate group is more unstable.

*Texts.* Some informal observations may be made on the texts. Figure 2 shows the mean word monitor latencies for each of the six pairs of experimental passages in the three subject groups.

As will be recalled, the target sentence was identical for the A and B versions in each of the six text pairs. Thus the monitor latencies for condition A and condition B were obtained, for each text pair, from the same word. In the native speaker group, four out of the six text pairs displayed clearly shorter monitor latencies for the A version as compared with the B version. A small difference in the expected direction was obtained for one pair (number 2), while one pair (number 1(, with 'Anthropologists' (A) and 'Neanderthal Man' (B) as topics, produced an advantage of some 70 msec for condition B. The target word (post-theme verb), which the subjects had to keep in mind presumably by rehearsal during the presentation of the target sentence, happened in this case to be a homophonous item (*passed*). For native speakers, this may have been a distracting factor that overrode the effect of the previous context and had an effect on monitoring performance through phonological coding and the semantic plausibility of the word *past* for both versions. One out of the nine native speaker subjects commented on the homophony. A likely explanation for the small A-B difference in text pair 2 is that the context included a third referent which in fact is fairly strongly in the foreground, especially in version A. This fact may have had an influence on topical expectations.

In the advanced non-native group, the texts function with a roughly similar degree of homogeneity. The deviant text pair is now pair 3, where the mean A latency is some 50 msec longer than the B latency. A relatively small difference is again observed for pair 2.

*Fig. 2.* Mean word monitor latencies for the six pairs of texts in the three subject groups.




In the INT group the A-B differences are generally small, with mean B latencies smaller than mean A latencies in three out of the six cases. Pair number 6, which was among the easiest lexically, produced a clear A < B difference even in this group, but otherwise the discourse-

topical effect observed with most texts in the NS and ADV groups did not appear in the INT group equally clearly.

### *3.6. Discussion*

Above, the effects of topical coherence on the comprehension processes of English native speakers and non-native (Finnish) users of English were observed in a simple two-factor experiment, with reaction times from a visual word monitoring task as the dependent measure. The word monitor task is a subsidiary task method (cf. Levelt 1978) that enables the investigator to focus on a specific point, selected on linguistic grounds, in order to estimate the effects of the linguistic variable on the on-going primary task of comprehension. The rationale for the technique is based on the assumption that spare capacity (cf. Dornic 1980, 1977a, 1977b) in the processing system can be directed to extraneous tasks. Performance in a subsidiary task reflects fluctuations in how available spare capacity is accessed; these fluctuations in turn are a function of the load imposed by the primary task. Another assumption is that the two simultaneous tasks are not processed totally separately by independent components of the processing system. In word monitoring, both the primary and secondary tasks involve linguistic operations. This seems to impose a higher load than using a nonverbal signal as the target, probably owing to greater structural interference (Dornic 1980) between the tasks. (An unpublished comparison by the present writer of phoneme and click monitoring latencies using isolated sentences and subject groups from the same level of proficiency suggests that monitoring for a nonverbal signal is faster by about a factor of two.) Structural interference is a problem if one wants to employ subsidiary techniques in comparisons of the demands imposed by different primary tasks. The subsidiary task cannot be used in any absolute or general sense to estimate the amount of capacity taken up or saved. However, for the present task, this interference would appear to ensure that latencies become sufficiently long and the relative differences in the effects of the independent linguistic variable are highlighted. The monitor technique seems to be able to tap on-going comprehension processes, including the predictive utilisation of contextual information. As it is a simple rather than a choice RT task, and involves rapid and immediate responses,

it also contains fewer conscious problem-solving elements than some other RT measures used in comprehension studies.

Distinct quantitative differences in word-monitor times appeared between the three proficiency groups. Non-native readers take considerably longer than native readers to react to verbal stimuli in the task, even when non-native proficiency is sufficiently high to allow seemingly effortless communication with native speakers. This suggests that the speed of the processes involved in the conscious detection of the target word from the stimulus display while simultaneously performing the primary task of general comprehension is dependent on proficiency level. A crucial factor would seem to be the variation in the subjects' ability to utilise regularities in the previous context, including the obvious syntactic predictabilities and the discourse-level dependencies associated with topical coherence. The facilitating effect of syntactic and other expectations on word monitoring was evident in the performance of those native speakers who occasionally responded on contextual grounds before the target word had appeared. (Some of the distractor tasks contained cases that induced strong syntactic expectations.) When the subject's proficiency level is high, the message can be effectively chunked in the sense of Miller (1967), capacity remains available for the subsidiary task, and monitor performance is fast. This interpretation is basically in line with observations of bilinguals, whose performance is markedly slower in the nondominant language in tasks such as acting out instructions, detecting sequences or combinations of numbers presented in the two languages, or matching words with pictures (Dornic 1977b).

The significant effect of context type resulted from the large A-B differences in the native speaker and advanced learner groups. Topical continuity appears to facilitate on-line processing in proficient readers. Having determined the discourse topic on the basis of surface cues, these readers seem to begin the current sentence expecting to encounter the topical referent again in the theme, unless, of course, a coming shift of topic is being signalled. If the sentence is in accordance with this predictive strategy, as in condition A, rapid referential mapping is possible, and spare capacity can be directed to the concurrent subsidiary task. Thus, processing load is relatively low at the locus of the expected discourse-topical referent and immediately following it. This also means that processing capacity can be efficiently used for the integration of the subsequent new information into existing memory struc-

tures. However, in condition B texts, the structure of the discourse does not conform to the predictive strategy. As a result, referential mapping across two sentences, with the reinstatement of the antecedent for the target sentence theme, introduces an increase in processing load, which is reflected in the word monitor latencies. It should be noted that the complicating effect of condition B remains below the threshold of consciousness. Subjects were interviewed briefly after the session, and none reported having noticed anything peculiar about any of the passages, though plenty of comments were offered on the comprehension questions and their relation to perceived gist. The latter fact was taken as a sign that instructions had successfully directed subjects' attention to the necessity of reading the materials for meaning.

The observed interaction between proficiency and topic is due to the performance of the intermediate non-native group, whose monitor latencies are not only quantitatively but also qualitatively different from native speaker and advanced non-native results, since there is little difference between the mean response times for condition A and condition B. This suggests that there may be some difference in the discourse processing strategies and capacity allocation principles of more proficient and less proficient non-native readers. With less proficient readers, high-level expectations from preceding context and such macrostructural features as the discourse topic may not always guide processing to the same extent as in the case of skilled foreign-language users.

Along the lines suggested above in section 1.3, a possible explanation may be sought in the limited capacity of the comprehension system. Processing capacity is taken up by those linguistic features of the message that are the most crucial ones for the reader's effort after meaning at that particular point. For less advanced readers, the crucial features may lie in the microstructure, and handling these local features in small units is costly. The result is that little processing may take place on the deeper levels of discourse macrostructure and overall passage meaning. The top-down strategy of using the identified discourse topic predictively as a conceptual peg on which new information can be hung is logically possible only if sufficient processing capacity is available on top of that required by the on-going microanalyses. Thus, global comprehension of foreign-language text comes to depend heavily on the automation of the lower-level processes. This is basically also the finding in L1 word-recognition research (cf. Vellutino 1980:156): poor

native-language readers are less efficient in utilising context to aid comprehension because the deficiencies in the lower-order skills cause short-term memory problems.

In addition to proficiency variations, other factors may also be involved. The quality of the text obviously interacts with proficiency, determining in part what the size of the unit of analysis will be, and how likely it will be that the non-native reader can engage in the processing of macrostructure. Passage 6 in the present materials may have been easy for the majority of subjects in the INT group, with the result that monitor times in condition B were clearly longer than in condition A. The performance of the foreign language reader may also reflect the practices of instruction. If analysis of the code is emphasised, and word-by-word processing thereby encouraged, the skills involved in the construction of macrostructure and the utilisation of global cues in comprehension may develop late. It would also seem that much depends on reading instruction in the native language; if mother-tongue instruction contains practice in these things, an effect should also be observable in foreign-language processing. The material one habitually reads may also be a factor in the strategies employed by adult foreign-language readers though in the present case the groups of subjects are not likely to differ greatly with regard to the need for extensive reading strategies. It was also assumed that general cognitive processing styles (the serialist-holist dichotomy, cf. Entwistle 1980) would be randomly distributed in the groups of subjects.

It is sometimes claimed in somewhat general terms that it is the less fluent non-native readers who are especially dependent on contextual factors in comprehension, since their microprocesses, notably word recognition, are slower and less efficient. This view is an extension into text comprehension of the compensatory view of L1 word recognition (e.g. Stanovich 1980). The essential claim of the compensatory view is that good L1 readers differ from poor readers in the speed of lexical recognition. The ability to utilise context, while present in both types of reader, does not necessarily differentiate between them. The crucial factor is the speed of context- free word recognition, which is so fast in good readers that reliance on contextual cues is seldom necessary. When contextual facilitation of word recognition does take place in these readers, it is through automatic spread-of-activation processes which require no cognitive capacity. Poor readers, on the other hand, need to compensate for their slow context-independent word recogni-

tion by the use of higher-order expectations that are slower, require cognitive capacity, and have time to operate precisely because context-independent word recognition is slow. Thus, for the process of lexical recognition, it may be said that while both good and poor readers can utilise context to aid word recognition, the good readers are usually less dependent on context than poor readers.

The extension of this view to comprehension of text and to the present case of the discourse topic would seem to predict that the intermediate readers should be particularly dependent on the variable of context type. The discourse topic, strongly emphasised by a simple title and consolidated through thematic iteration, is a semantic element that could be expected to stand out for the intermediate readers in material that otherwise may remain more or less vague. The prediction would thus be that the intermediate readers should evidence contextual facilitation (condition A) and inhibition (condition B) in their word monitor times at least as clearly as the advanced non-natives and native speaker subjects. A difference would be expected in the speed of word recognition, which should be slower for the intermediate group. Insofar as word monitor times can be used as an index of the speed of lexical recognition, such a difference was obtained. Otherwise, the results appeared to suggest that it is the advanced rather than the intermediate non-native readers who tend to be consistently sensitive to higher-order aspects of the message (the discourse topic), and tend to use it to guide the on-line comprehension of the text.

The compensatory hypothesis thus does not necessarily extend to the integrative and predictive processes of understanding. The distinction between word recognition and text understanding is in fact highlighted by consideration of the consequences of compensatory word recognition for poor readers (cf. Stanovich 1980:64). Using the attention-demanding mechanisms of expectation to aid word recognition (i.e., having to utilise clues from, say, syntax to make up for the fact that the word cannot automatically be recognised as a unit) reduces the capacity that is available for the higher-order processes of comprehension, whereby a meaning representation is built for the sentences, and, on the level of macrostructure, for entire texts. Much of the processing capacity of the poorer readers (less proficient foreign-language readers) thus goes into various microprocesses, which may impede the handling and predictive utilisation of such macrofeatures as the discourse topic, gist, the writer's attitude, and nonliteral meaning.

The experiment described above attempted to tap some of the higher-order processes in comprehension at a specific point during processing. What was being measured during the target sentence was the fluctuation in processing load at the point when the thematically interesting criterial word (*Fry* in Table 1) has been recognised and the reader is processing the next word (the target verb) while simultaneously performing the conscious vigilance task of word monitoring. The monitor latency to the target verb can be taken to reflect load from processes whereby the reader is attempting to relate the criterial word just recognised to the information derived from the preceding discourse. For advanced non-native readers, whose processing capacity is not heavily taxed by microprocesses, the facilitatory and inhibitory effects of topical structure on the integration processes are clear. For less advanced foreign language readers, whose comprehension of the topic in the preceding context and its top-down utilisation is probably less efficient, topical discontinuity in the preceding context does not have a similar consistent effect.

## REFERENCES


Britton, B.K., Piha, A., Davis, J. and Wehausen, E. 1978. Reading and cognitive capacity usage: adjunct question effects. *Memory & Cognition* 6, 266-273.

Britton, B.K., Meyer, B.J.F., Simpson, R., Holdredge, T.S., and Curry, C. 1979. Effects of the organization of text on memory: Tests of two implications of a selective attention hypothesis. *Journal of Experimental Psychology* 5, 496-506.

Chafe, W.L. 1972. Discourse structure and human knowledge. In Freedle, R. and Carroll, J.B. (eds.), *Language Comprehension and the Acquisition of Knowledge*. Washington: Winston.

Chihara, T., Oller, J. Weaver, K. and Chavez-Oller, M.A. 1977. Are cloze items sensitive to constraints across sentences? *Language Learning* 27, 63-73.

Clark, H.H. and Haviland, S.E. 1977. Comprehension and the given-new contract. In Freedle, R.O. (ed.), *Discourse Production and Comprehension*. Norwood, N.J.: Ablex.

Clarke, M.A. and Silberstein, S. 1977. Toward a realization of psycholinguistic principles in the ESL reading class. *Language Learning* 27, 135-154.

Clements, P. 1979. The effects of staging on recall from prose. In Freedle, R. (ed.), *New Directions in Discourse Processing*. Norwood, N.J.: Ablex.

Cohen, A. and Hosenfeld, C. 1981. Some uses of mentalistic data in second language research. *Language Learning* 31, 285-313.

Cziko, G.A. 1977. Difference in first- and second-language reading: the use of semantic and discourse constraints. *The Canadian Modern Language Review* 34, 473-489.

Cziko, G.A. 1980. Language competence and reading strategies: a comparison of first- and second-language oral reading errors. *Language Learning* 30, 101-116.

van Dijk, T.A. 1977. Sentence topic and discourse topic. *Papers in Slavic Philology* 1, 49-61.

van Dijk, T.A. 1980. *Macrostructures*. Hillsdale, N.J.: Lawrence Erlbaum.

Dooling, D.J. and Lachman, R. 1971. Effects of comprehension on retention of prose. *Journal of Experimental Psychology* 88, 216-222.

Dornic, S. 1977a. *Mental load, effort, and individual differences*. Reports from the Department of Psychology, University of Stockholm, N:o 509.

Dornic, S. 1977b. *Information processing and bilingualism*. Reports from the Department of Psychology. University of Stockholm, N:o 510.

Dornic, S. 1980. *Spare capacity and perceived effort in information processing*. Reports from the Department of Psychology, University of Stockholm, N:o 567.

Enkvist, N.E. 1973. 'Theme dynamics' and style: an experiment. *Studia Anglica Posnaniensia* 5, 127-135.

Enkvist, N.E. 1975. *Tekstilingvistiikan peruskäsitteitä*. Helsinki: Gaudeamus.

Enkvist, N.E. 1980. Rapporteur's survey in the Section on Text Analysis and Generation. Paper read at the Nobel Conference on Text Processing, Stockholm 1980.

Entwistle, N. 1981. *Styles of Learning and Teaching*. Chichester etc.: John Wiley.

Frederiksen, C.H. 1977. Semantic processing units in understanding text. In Freedle, R.O. (ed.), *Discourse Production and Comprehension*. Norwood, N.J.: Ablex.

Godfrey, D. 1977. Listening instruction and practice for advanced second language students. *Language Learning* 27, 109-122.

Grimes, J.E. 1975. *The Thread of Discourse*. The Hague: Mouton.

Hosenfeld, C. 1977. A preliminary investigation of reading strategies of successful and non-successful second language learners. *System* 5, 110-123.

Jacobs, S.E. 1981. Rhetorical information as predication. *TESOL Quarterly* 15, 237-249.

Keenan, E. Ochs and Schieffelin, B.B. 1975. Topic as a discourse notion: a study of topic in the conversations of children and adults. In Li, C. (ed.), *Subject and Topic*. New York etc: Academic Press.

Kieras, D.E. 1978. Good and bad structure in simple paragraphs: effects on apparent theme, reading time, and recall. *Journal of Verbal Learning and Verbal Behavior* 17, 13-28.

Kieras, D.E. 1981a. Component processes in the comprehension of simple prose. *Journal of Verbal Learning and Verbal Behavior* 20, 1-23.

Kieras, D.E. 1981b. The role of major referents and sentence topics in the construction of passage macrostructure. *Discourse Processes* 4, 1-15.

Kintsch, W. and van Dijk, T.A. 1978. Toward a model of text comprehension and production. *Psychological Review* 85, 363-393.

Lautamatti, L. 1980. Subject and theme in English discourse. In Sajavaara, K. and Lehtonen, J. (eds.), *Papers in Discourse and Contrastive Discourse Analysis.* Jyväskylä Contrastive Studies 5.

Lesgold, A.M., Roth, S.F. and Curtis, M.E. 1979. Foregrounding effects in discourse comprehension. *Journal of Verbal Learning and Verbal Behavior* 18, 291-308.

Lesgold, A.M. and Perfetti, C.A. 1978. Interactive processes in reading comprehension. *Discourse processes* 1, 323-336.

Levelt, W.J.M. 1978. A survey of studies in sentence perception. In Levelt, W.J.M. and Flores d'Arcais, G.B. (eds.), *Studies in the Perception of Language.* Chichester etc.: John Wiley.

Meyer, B.J.F. 1975. *The Organisation of Prose and its Effects on Memory.* Amsterdam etc.: North-Holland.

Miller, G.A. 1967. *The Psychology of Communication: Seven Essays.* Harmondsworth: Penguin.

Perfetti, C.A. and Goldman, S.R. 1974. Thematization and sentence retrieval. *Journal of Verbal Learning and Verbal Behavior* 13, 70-79.

Perfetti,C.A. and Goldman, S.R. 1975. Discourse functions of thematization and topicalization. *Journal of Psycholinguistic Research* 4, 257-271.

Salmela, J. and Tommola, J. 1982. Apuvälineitä psykolingvistiseen tutkimukseen. *Korkeakoulujen ATK-uutiset* 3/1982, 21-22.

Sanford, A.J. and Garrod, S.C. 1981. *Understanding Written Language.* Chichester etc.: John Wiley.

Stanovich, K.E. 1980. Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly* XVI, 32-71.

Tannen, D. 1979. What's in a frame? Surface evidence for underlying expectations. In Freedle, R.O. (ed.), *New Directions in Discourse Processing.* Norwood. N.J.: Ablex.

Thorndyke, P.W. 1977. Cognitive Structures in comprehension and memory of narrative discourse. *Cognitive Psychology* 9, 77-110.

Tommola, J. Varantola, H. Pöllä, J. and Salmela, J. 1982. Välineistö puhutun ja kirjoitetun kielen ymmärtämistä koskeviin reaktioaikatutkimuksiin. In Sajavaara, K., Kalin, M. and Leiwo, M. (eds.), *Psykolingvistisia kirjoituksia III.* Jyväskylä: AFinLA.

Tommola, J. Varantola, H., Pöllä, J. and Salmela, J. (forthcoming). A reaction time analyser system for investigating the comprehension of spoken and written language.

Vellutino, F.R. 1982. Theoretical issues in the study of word recognition: The unit of perception controversy reexamined. In Rosenberg, S. (ed.), *Handbook of Applied Psycholinguistics*. Hillsdale, N.J.: Lawrence Erlbaum.